Trends in Reanalysis Data

Weather and Climate Impact Assessment

by Eric Gilleland and Matt Pocernich

National Center for Atmospheric Research

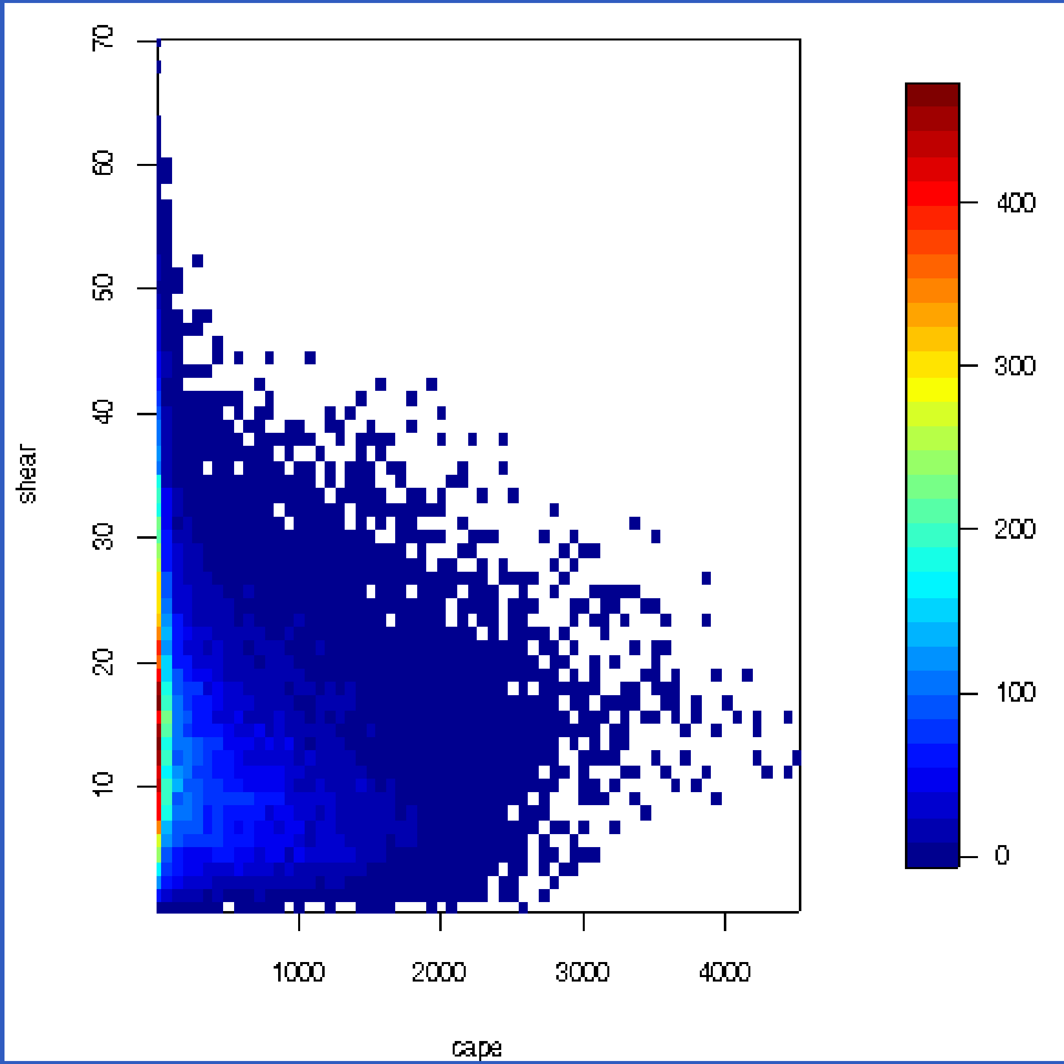
Outline

- Relations between CAPE and Shear and severe storms
- Reanalysis data
- Hypothesis tests and false discovery rates
- Simple linear trends
- Extreme value statistics
- Problems and future steps
- Comments, questions and suggestions.

CAPE, shear and severe storms

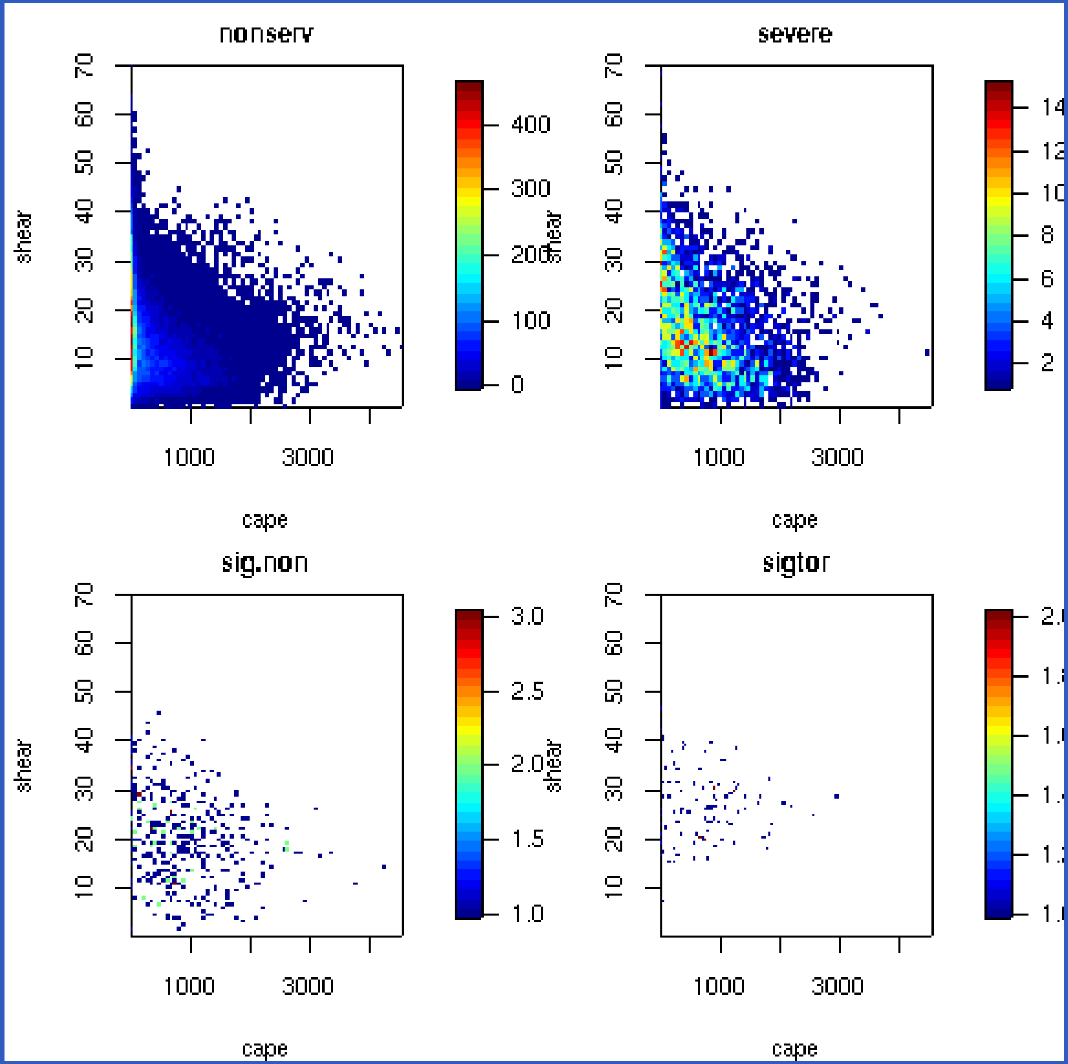
- CAPE = convective available potential energy; Shear = vertical velocity
- In order for a convective storm to exist, high amounts of both need to be present, though high quantities of both tend not to exist at the same time.
- Data with both observations and reanalysis data is relatively limited.
- Categorizes all events as either non-severe, severe, significant non-tornadic, severe-tornadic.

CAPE vs. Shear

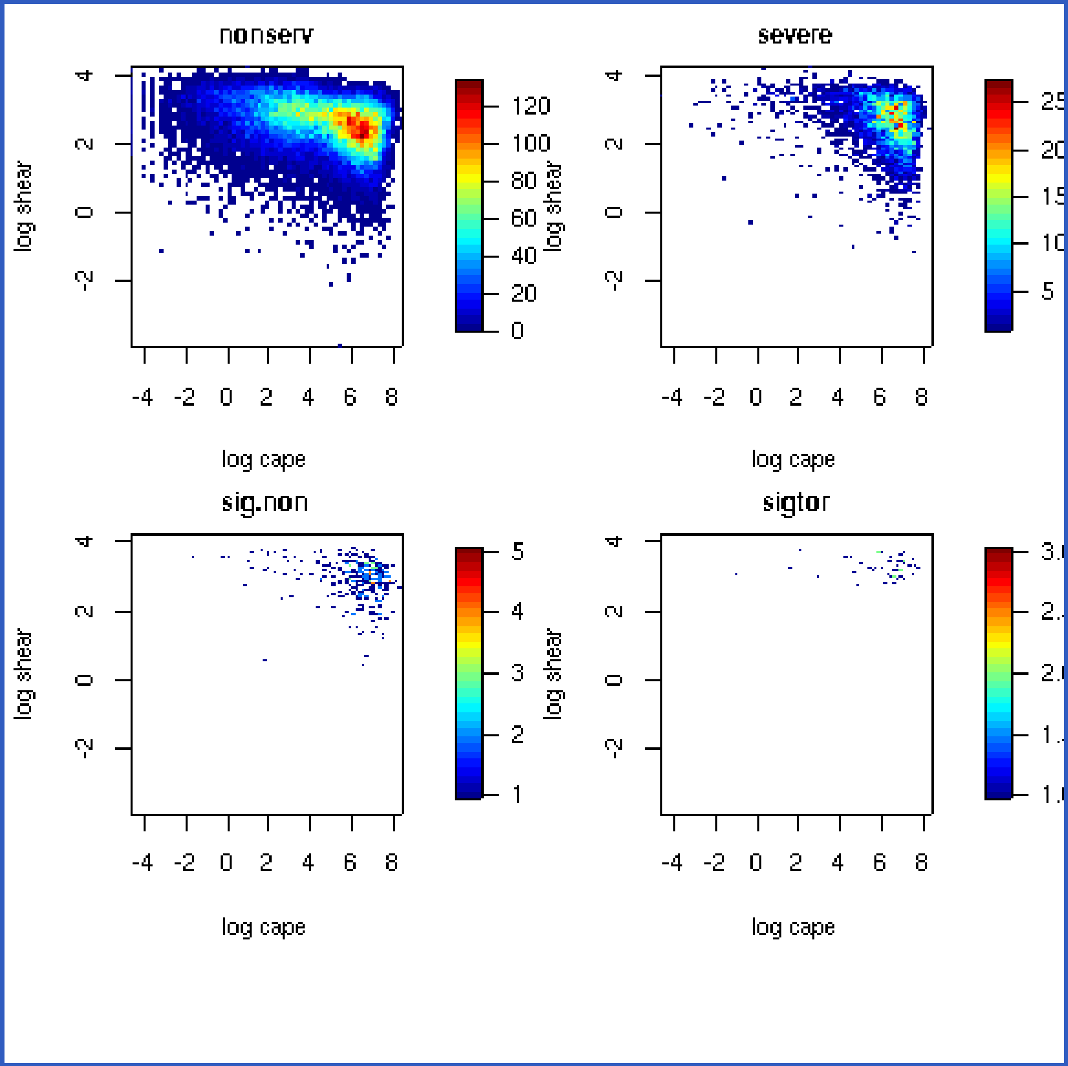


Select data: 1997 - 1999

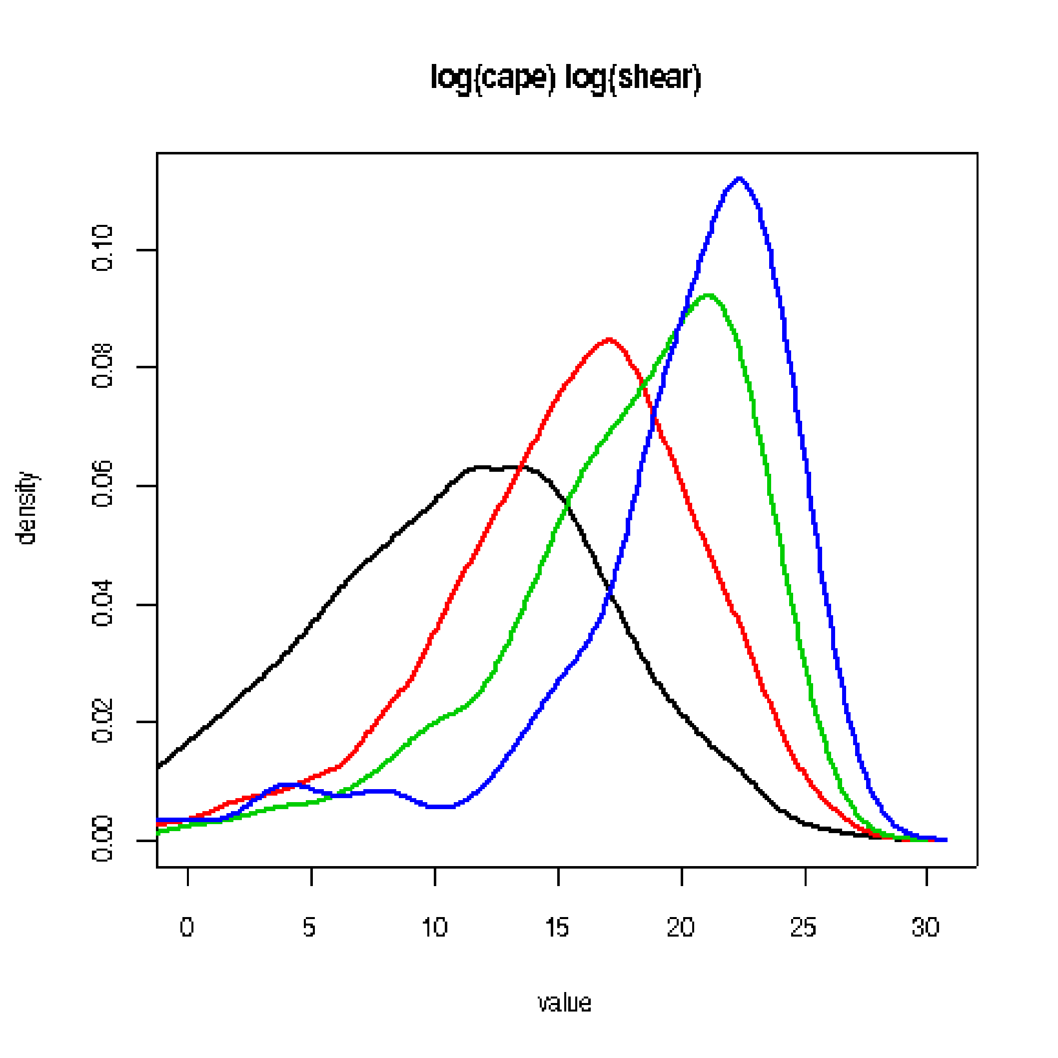
CAPE vs. Shear



Log CAPE vs. log Shear



log(CAPE)*log(Shear) Density Plots

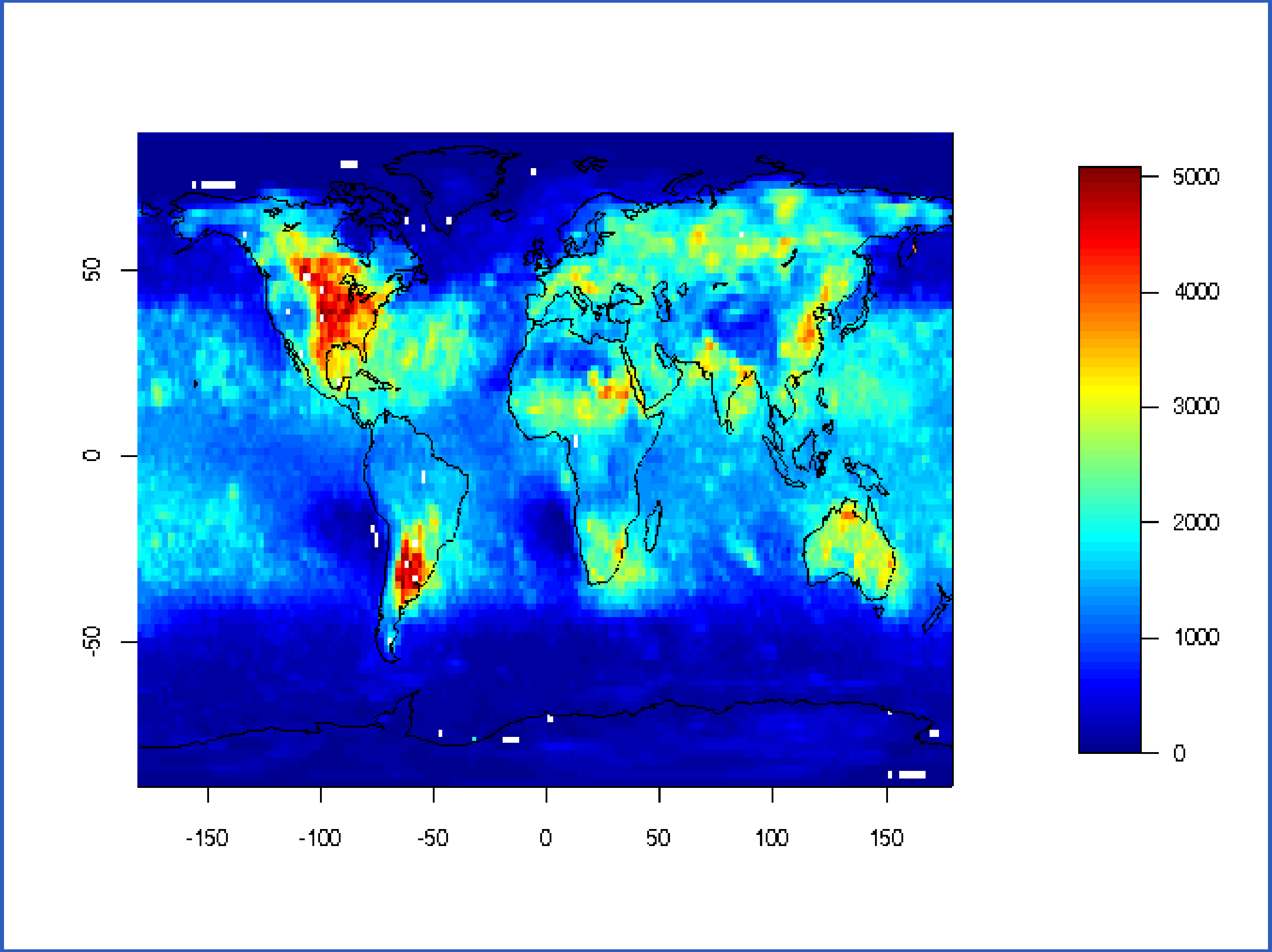


Select data: 1997 - 1999

Reanalysis Data (1957 - 1999)

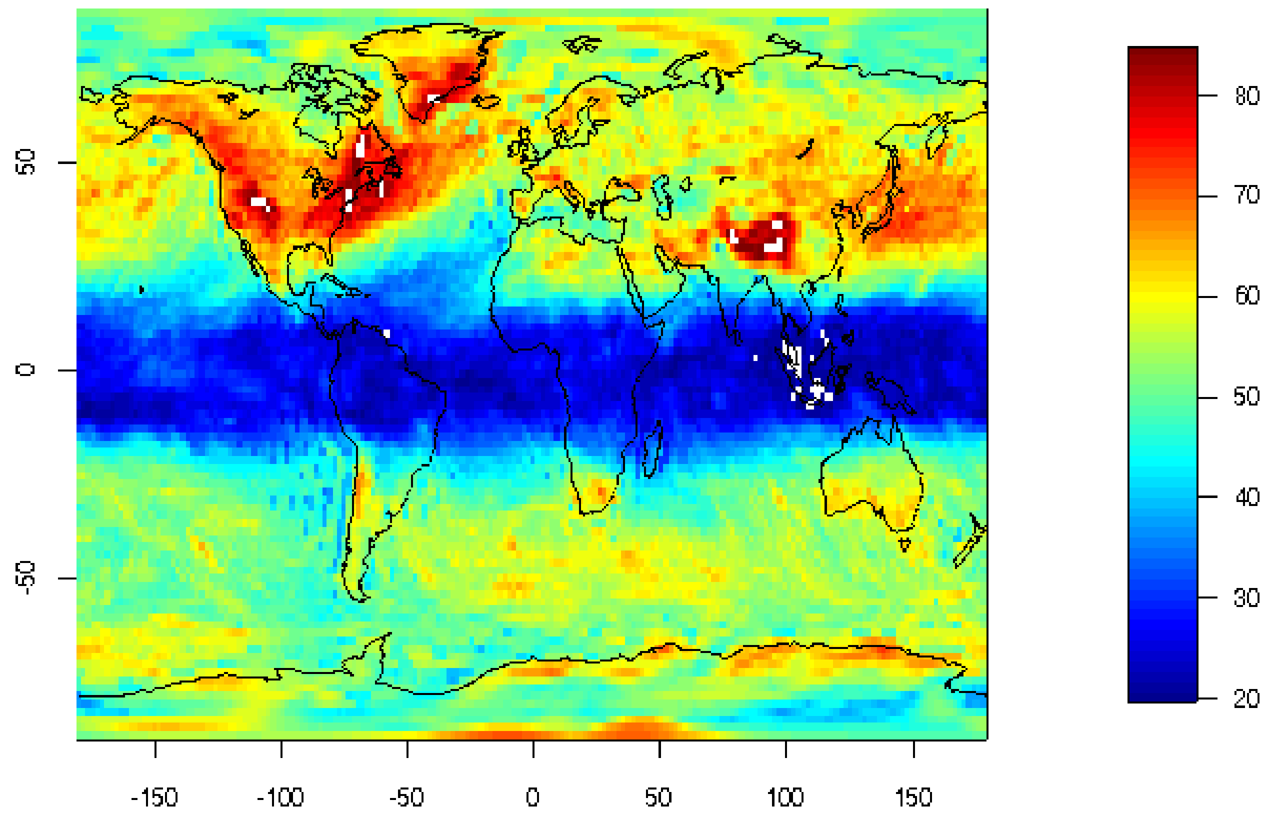
- Resolution 1.9x1.9 degree lat lon, 27 levels (Approx 17,000 gridpoints).
- Every 6 hours from June 1, 1957- Dec 31, 1999 with global coverage. (Currently we are using data from 1970 - 1999).
- Process very similar to starting up a global weather forecast model.
- Soundings were extracted from the data and run through the Sounding and Hodograph Analysis and Research Program (SHARP) from the NWS Storm Prediction Center (John Hart wrote the code).

Maximum Annual CAPE



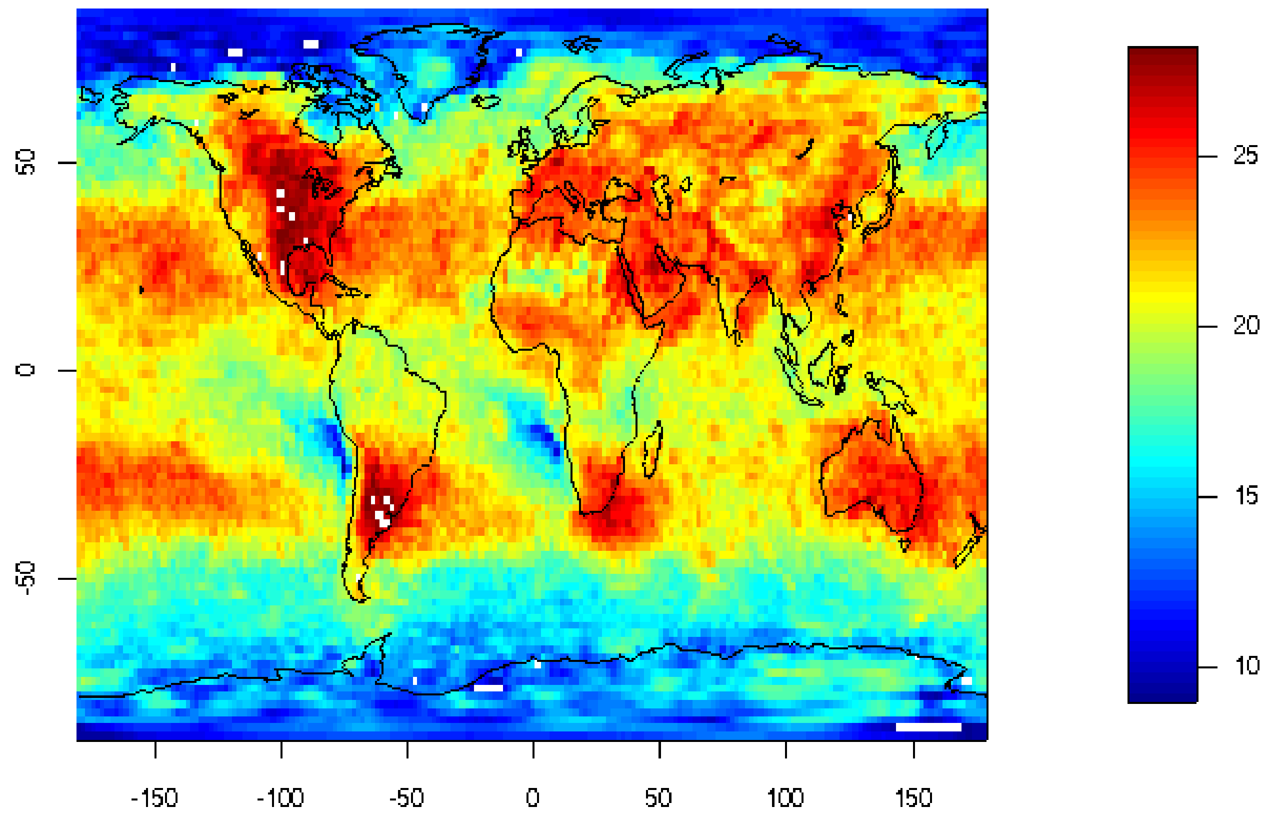
Data:1970-1999

Maximum Annual Shear



Data:1970-1999

Max Annual log(CAPE)*log(Shear)



Hypothesis Testing

Suppose n tests are performed, and a single test has probability α of rejecting the null hypothesis when it is, in fact, true. Then,

If all n null hypotheses are true, then on average, $n\alpha$ of them will be falsely rejected.

So, for example, if $n = 1000$ locations (and the null hypothesis is true for all 1000), then one expects 50 such locations will be incorrectly found to be significant.

Hypothesis Testing (con't)

Several methods are available to account for this problem. Two main themes:

field significance: Is an entire *field* significant?

multiple testing: Which *tests* are significant?

Hypothesis Testing (con't)

The first method does not give an indication of which particular locations are significant, but is very popular in climate studies (over 300 citations of Livezey and Chen (1983)).

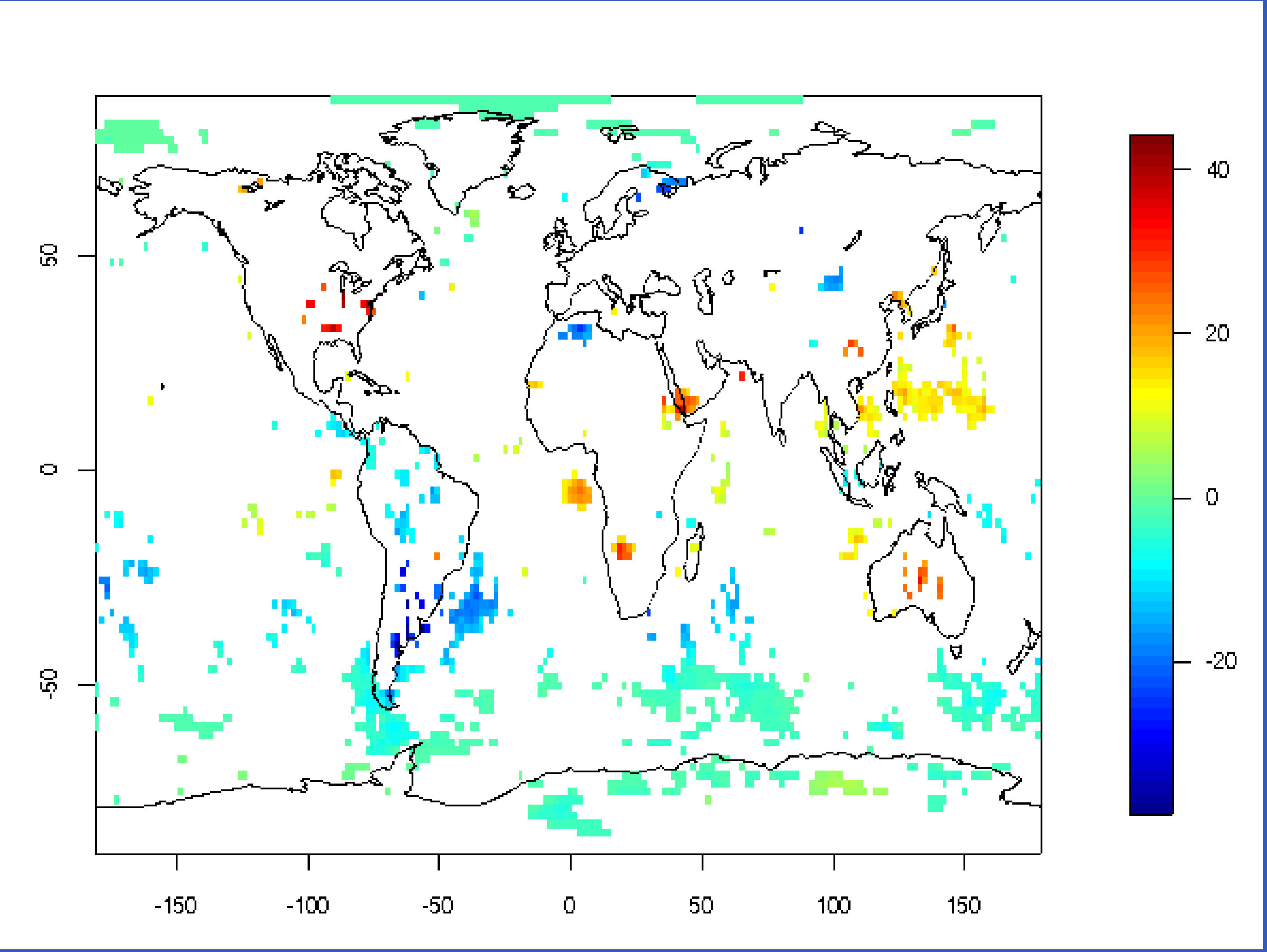
The second method makes local determinations of significance.

False Discovery Rate (FDR)

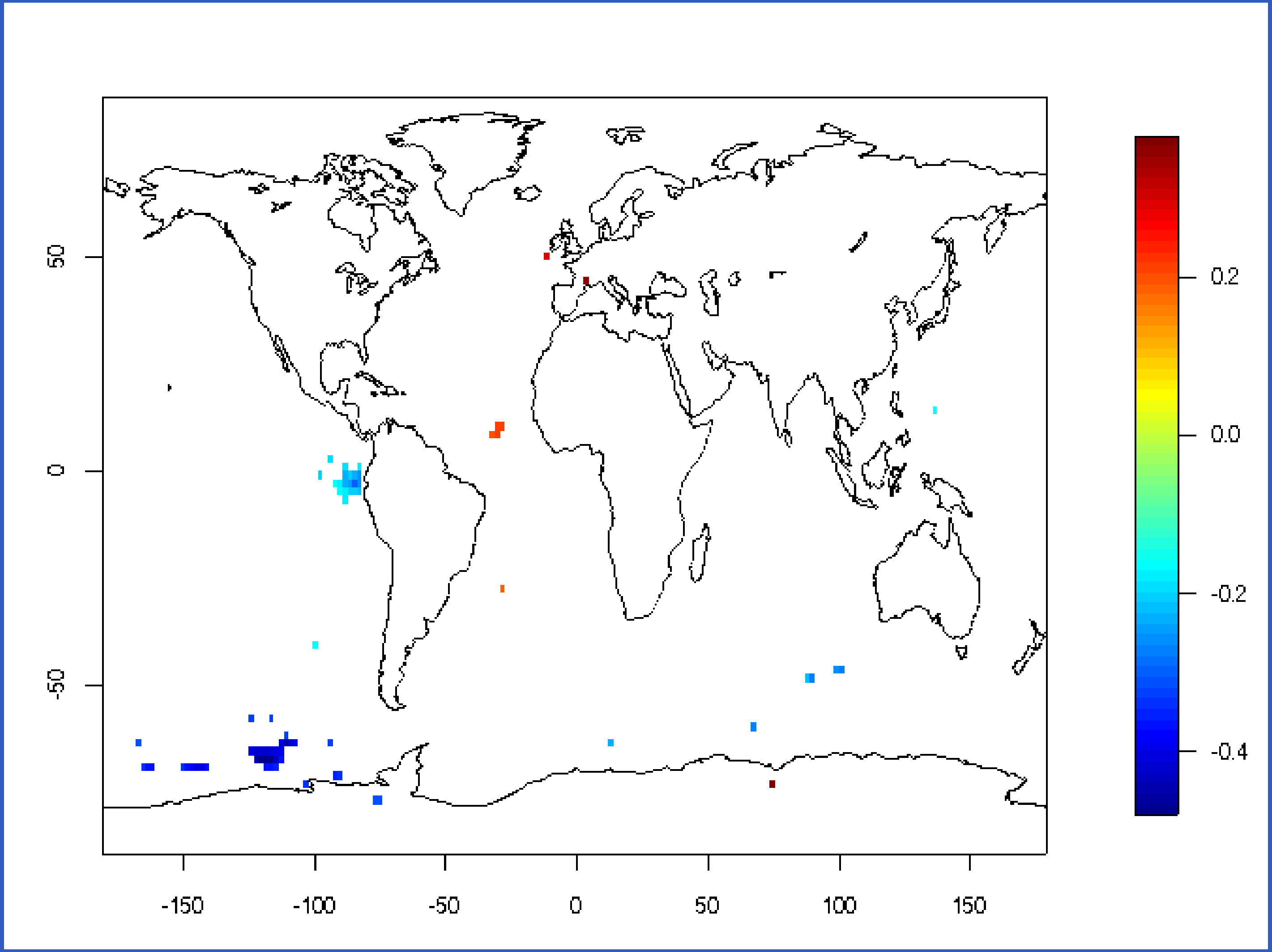
Typically, one controls a priori for the probability of detecting significance that does not exist.

Ventura et al (2004) argue for controlling for probability of falsely rejecting the null hypothesis for climatological studies instead. That is, probability falsely rejecting the null hypothesis. Also known as the false discovery rate (FDR). Controls the proportion of false rejections out of all rejections.

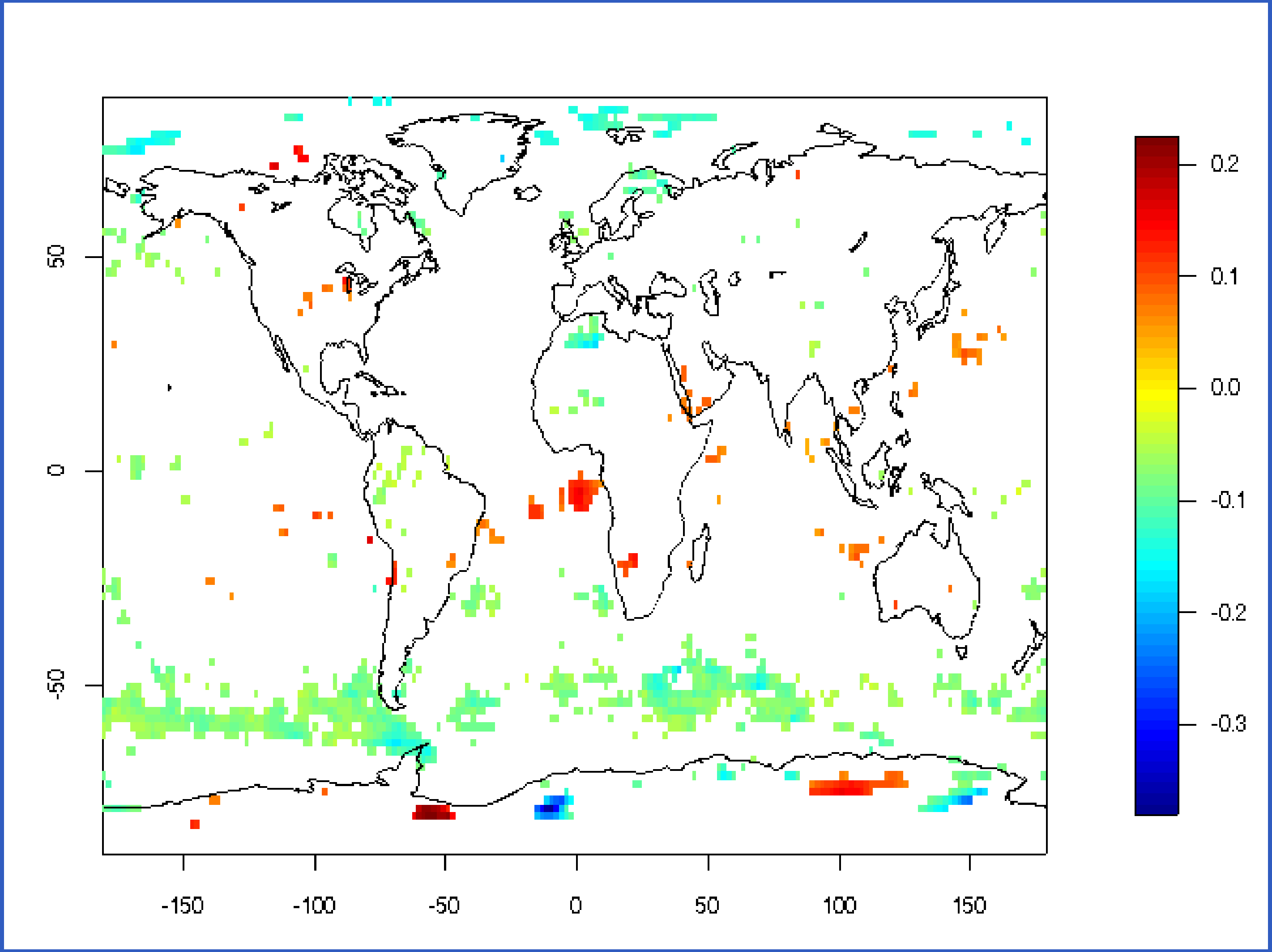
Max Annual CAPE



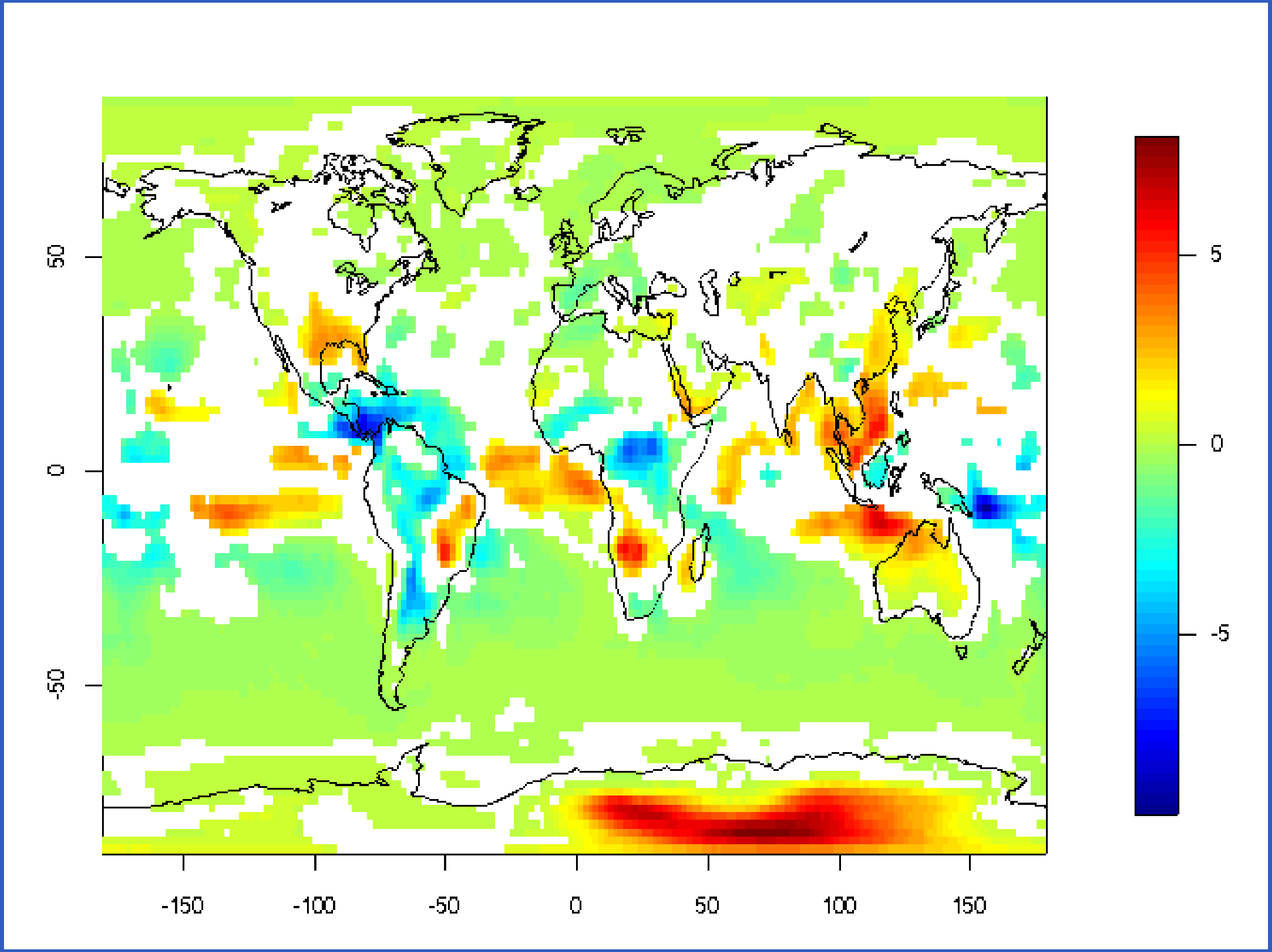
Max Annual Shear



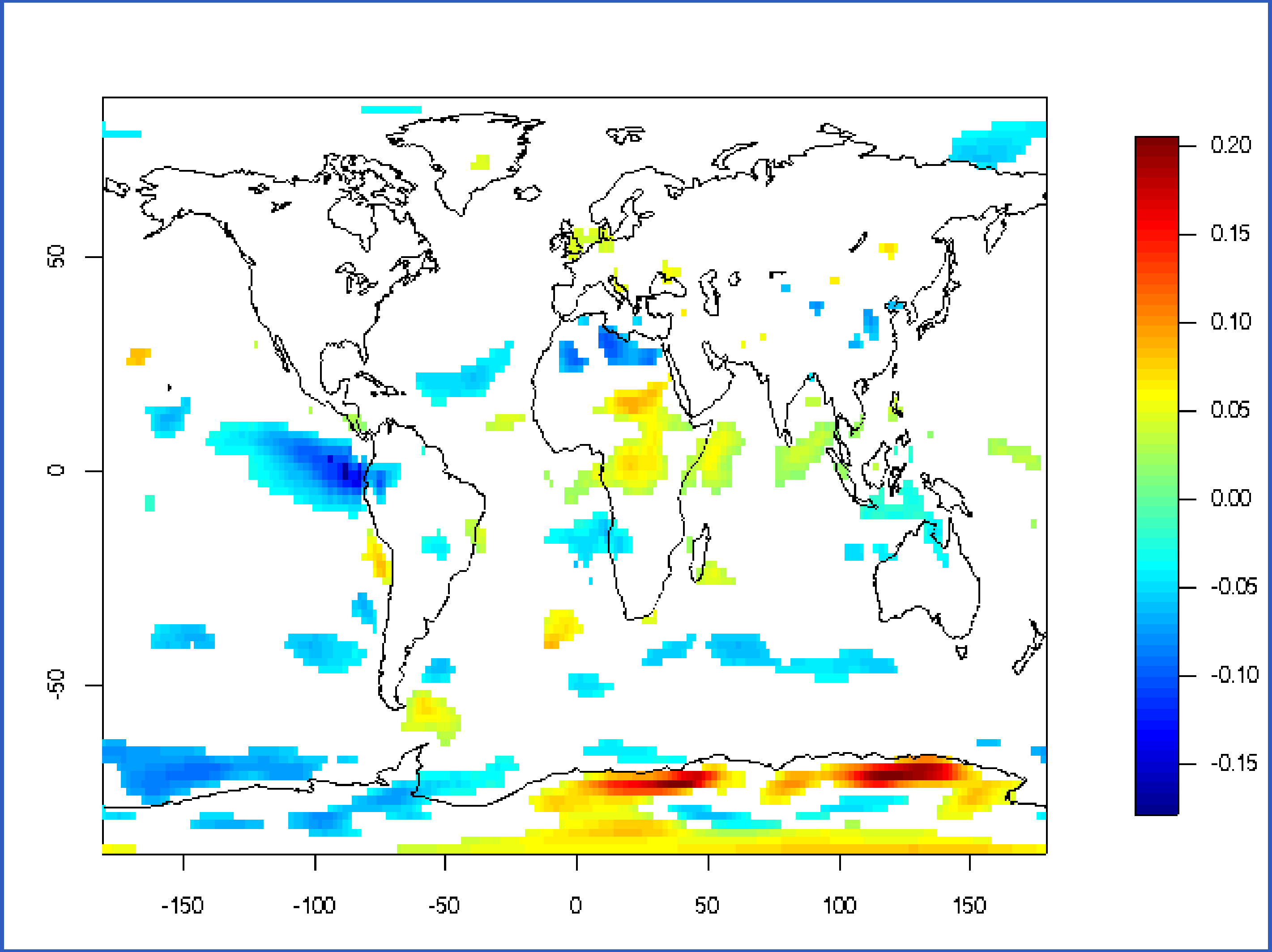
Max Ann. log(CAPE)log(Shear)



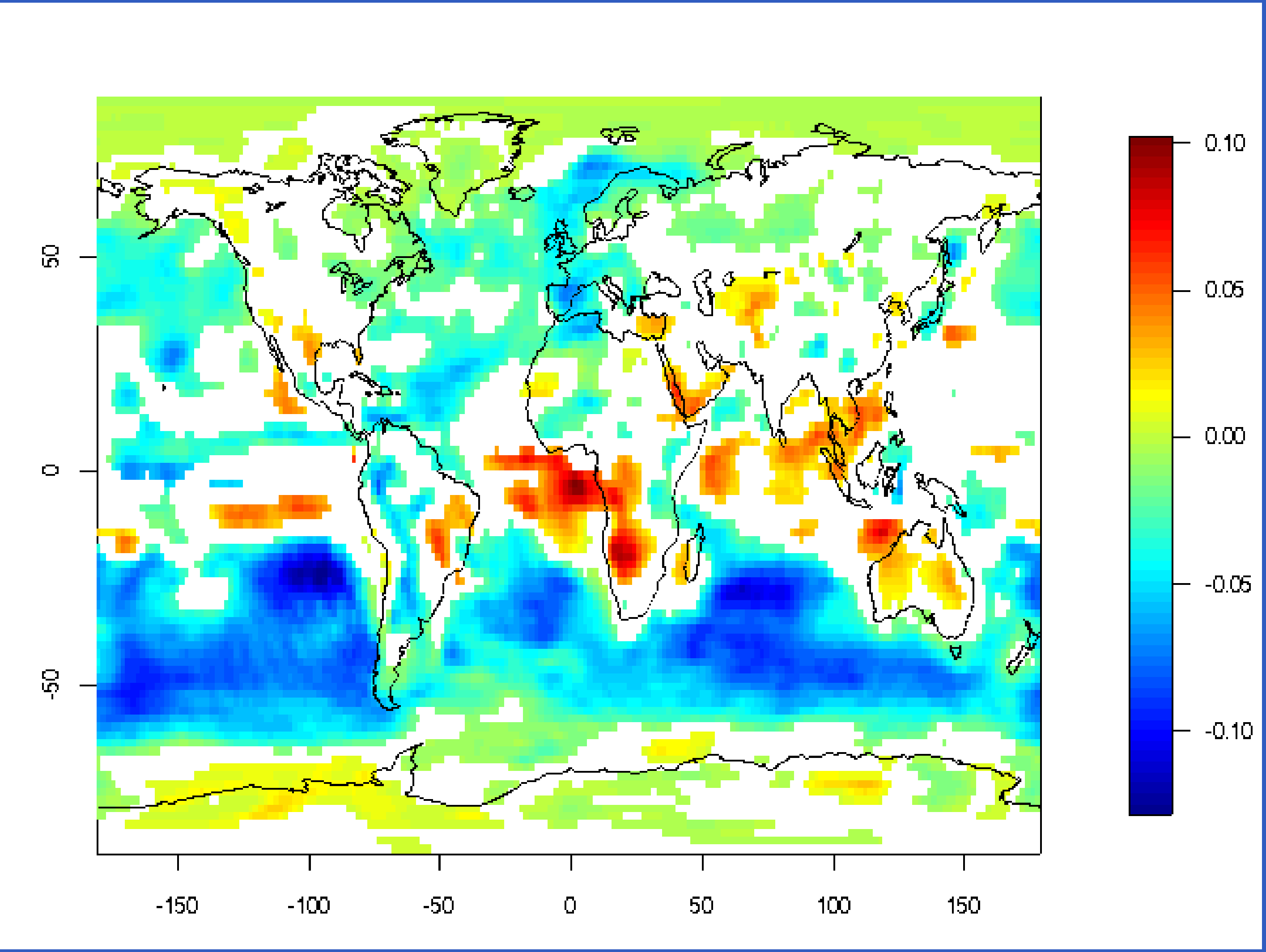
Mean Annual CAPE



Mean Annual Shear



Mean Ann. log(CAPE)log(Shear)



Fitting these data to the GEV

For now, fits are carried out at each grid point without accounting for any spatial structure.

Presently only 30 years of data, so MLE can be unstable, and often unable to fit using standard MLE method. Could use L-moments, but here we use a constrained MLE approach.

Constrained MLE

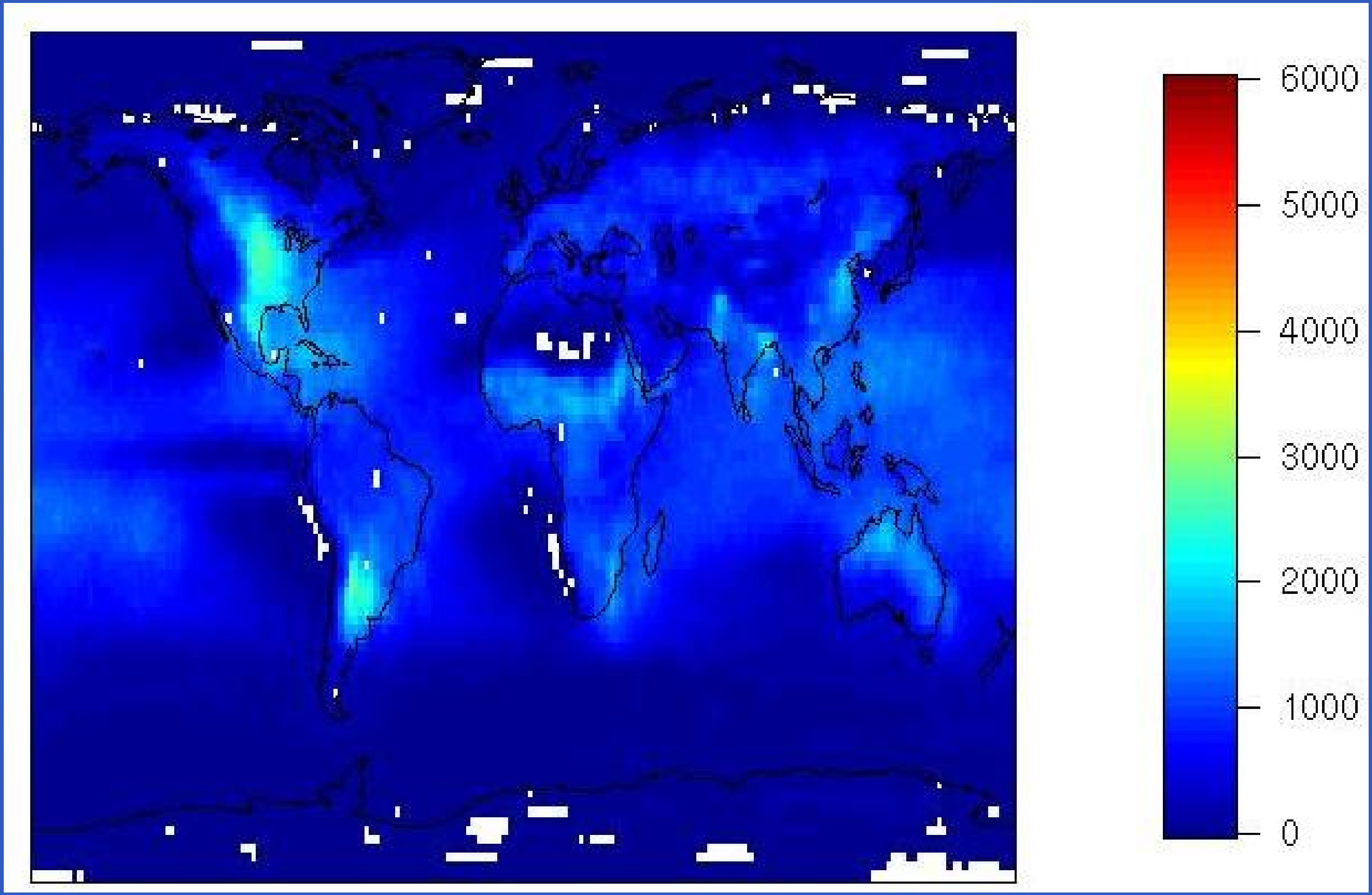
Approach is similar to that of Coles and Dixon (1999) and Martins and Stedinger (2000) (both papers have been discussed in this group).

However, no a priori information as to what ξ should be. So, we choose a relatively uninformed prior, $\pi(\xi)$, in (4.2) of Coles and Dixon.

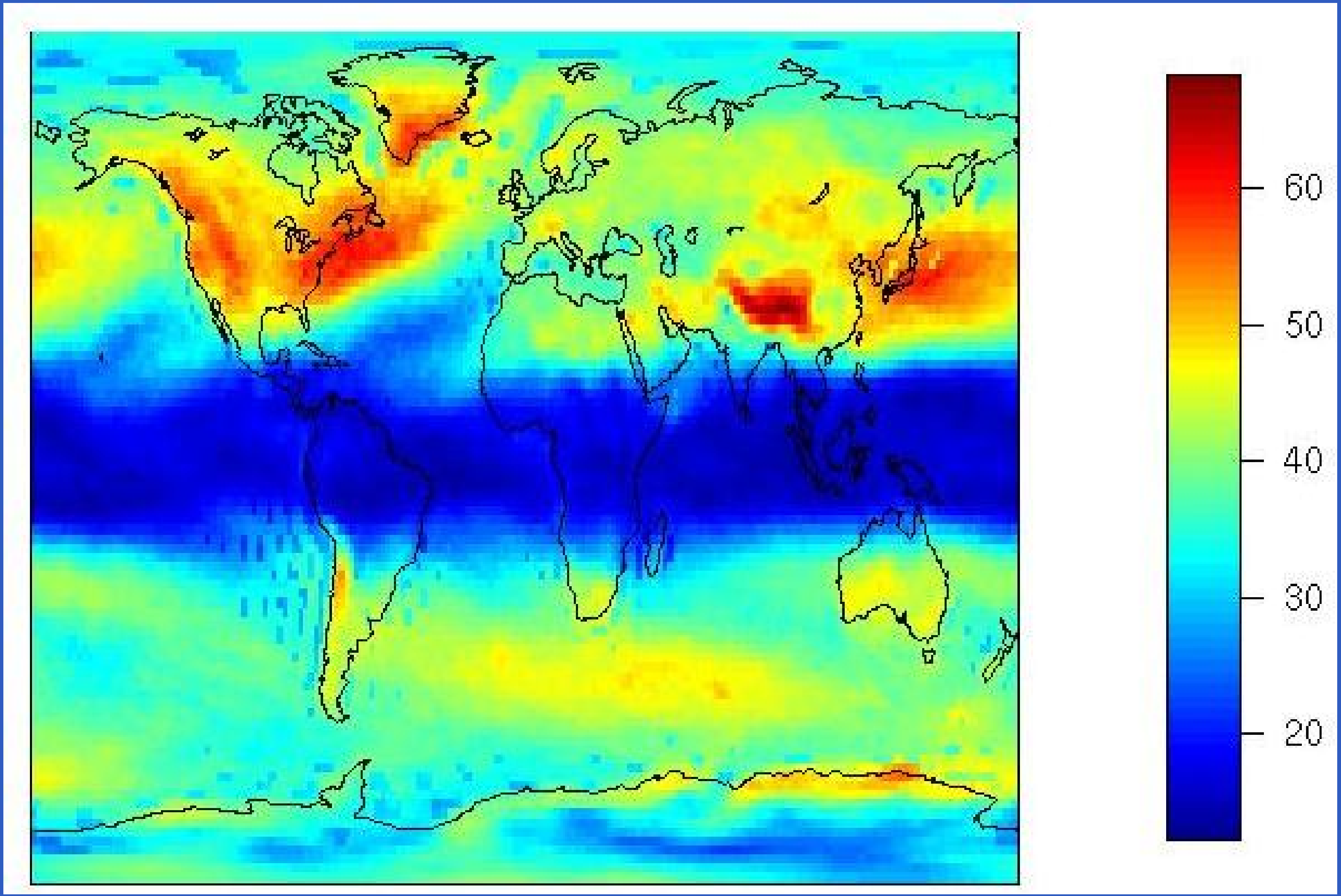
$$\ell_{\text{pen}}(\mu, \sigma, \xi) = \ell_{\text{GEV}}(\mu, \sigma, \xi) \times \pi(\xi),$$

where we use $\pi(\xi) = 1_{-0.5 \leq \xi \leq 1}$ (i.e., a uniform from -0.5 to 1). We don't trust models fit at grid points where ξ is close to boundaries, but this did not happen at most gridpoints (about 3% for ann max cape and ann max shear, and about 9% for log(cape)log(shear)).

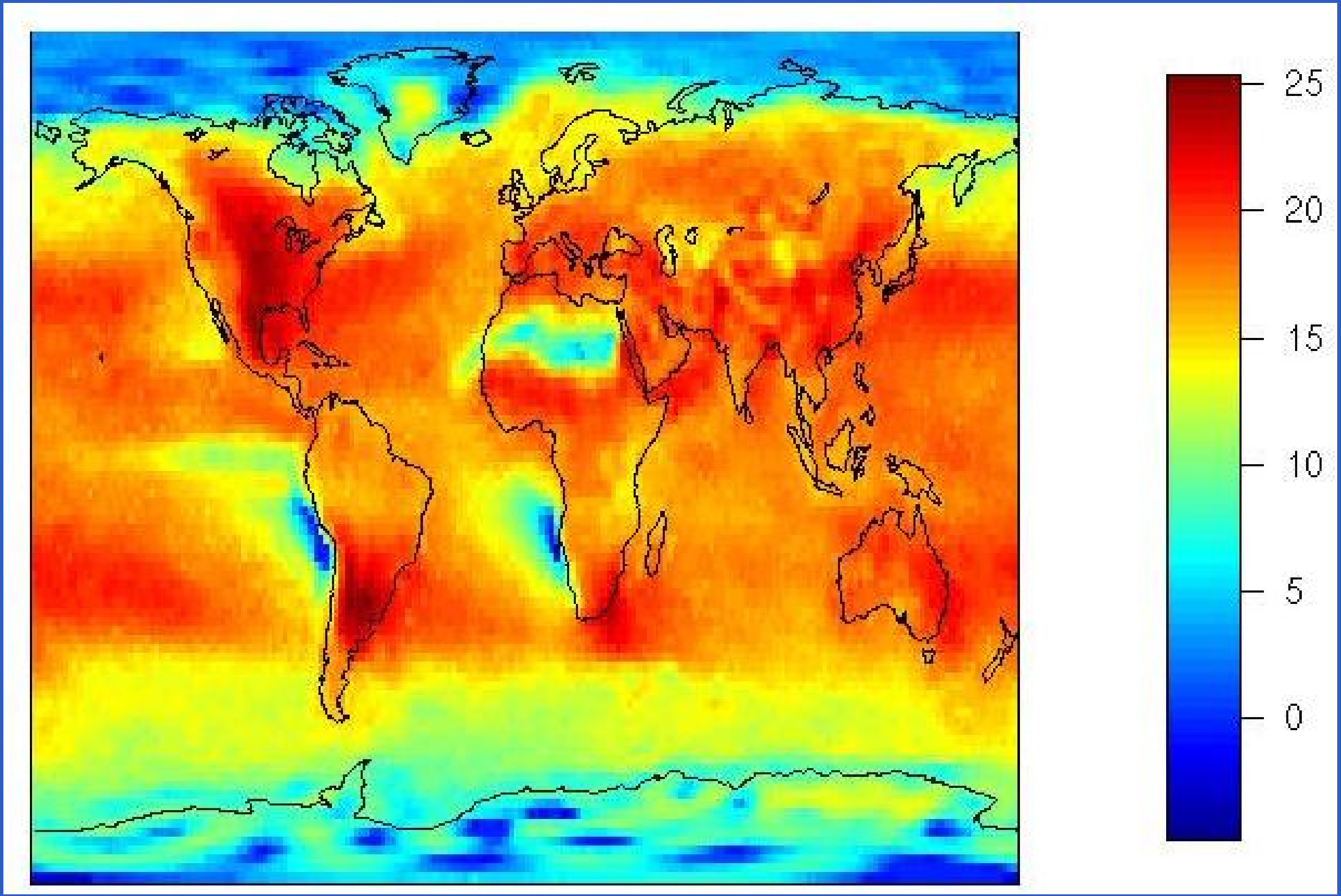
20 Yr Return Level - Max CAPE



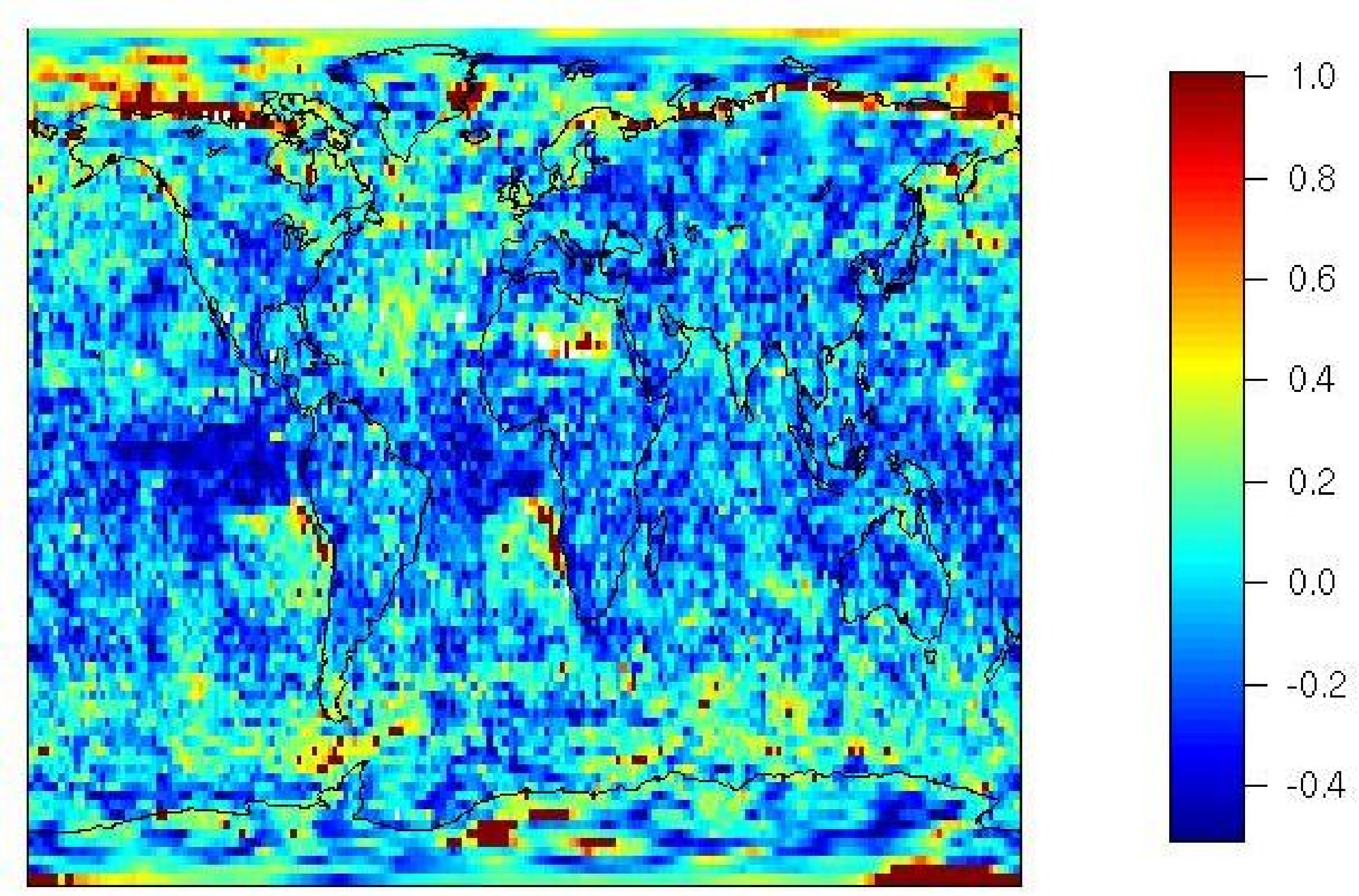
20 Yr Return Level - Max Shear



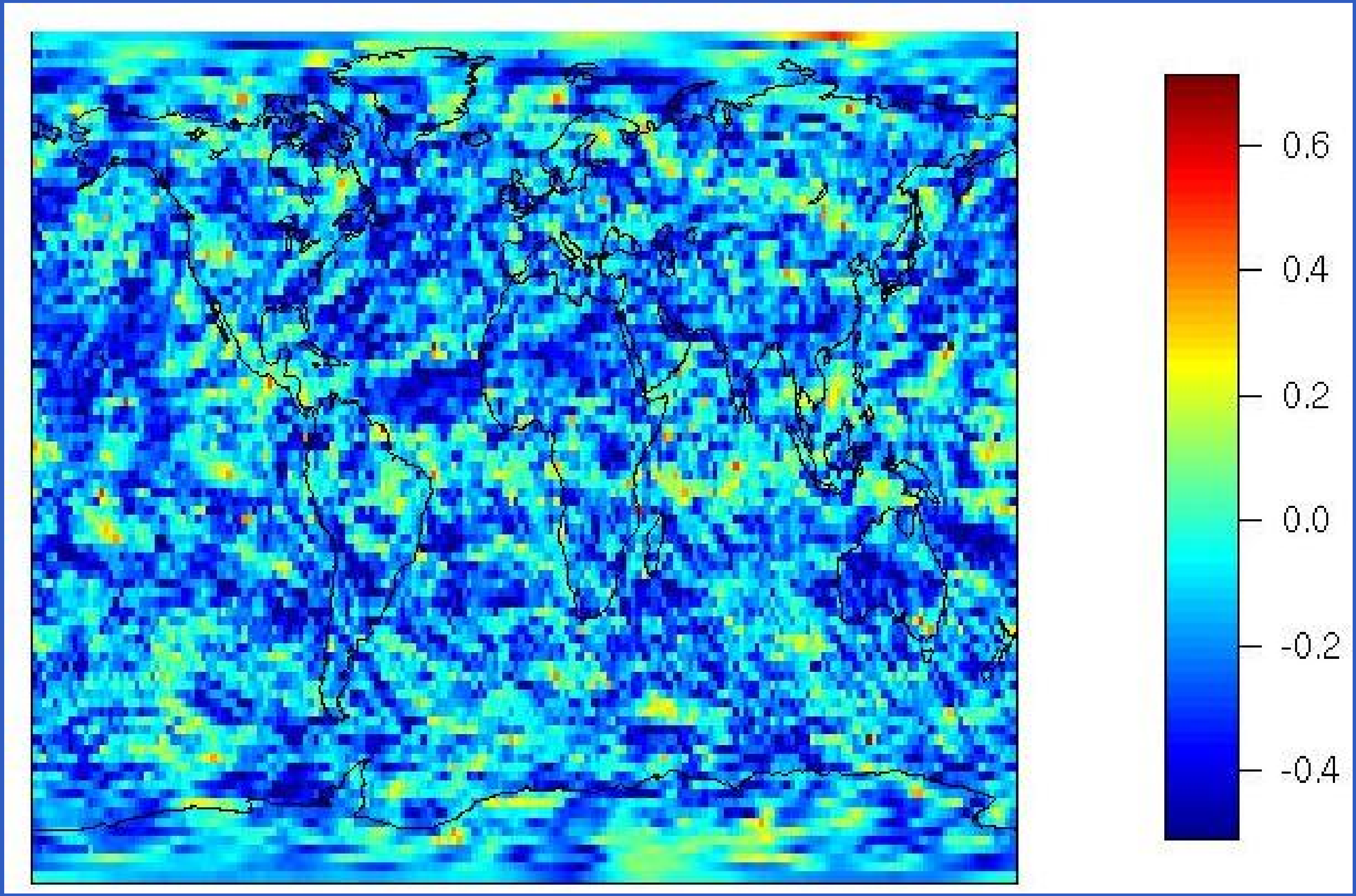
20 Yr Return Level - Max log*log



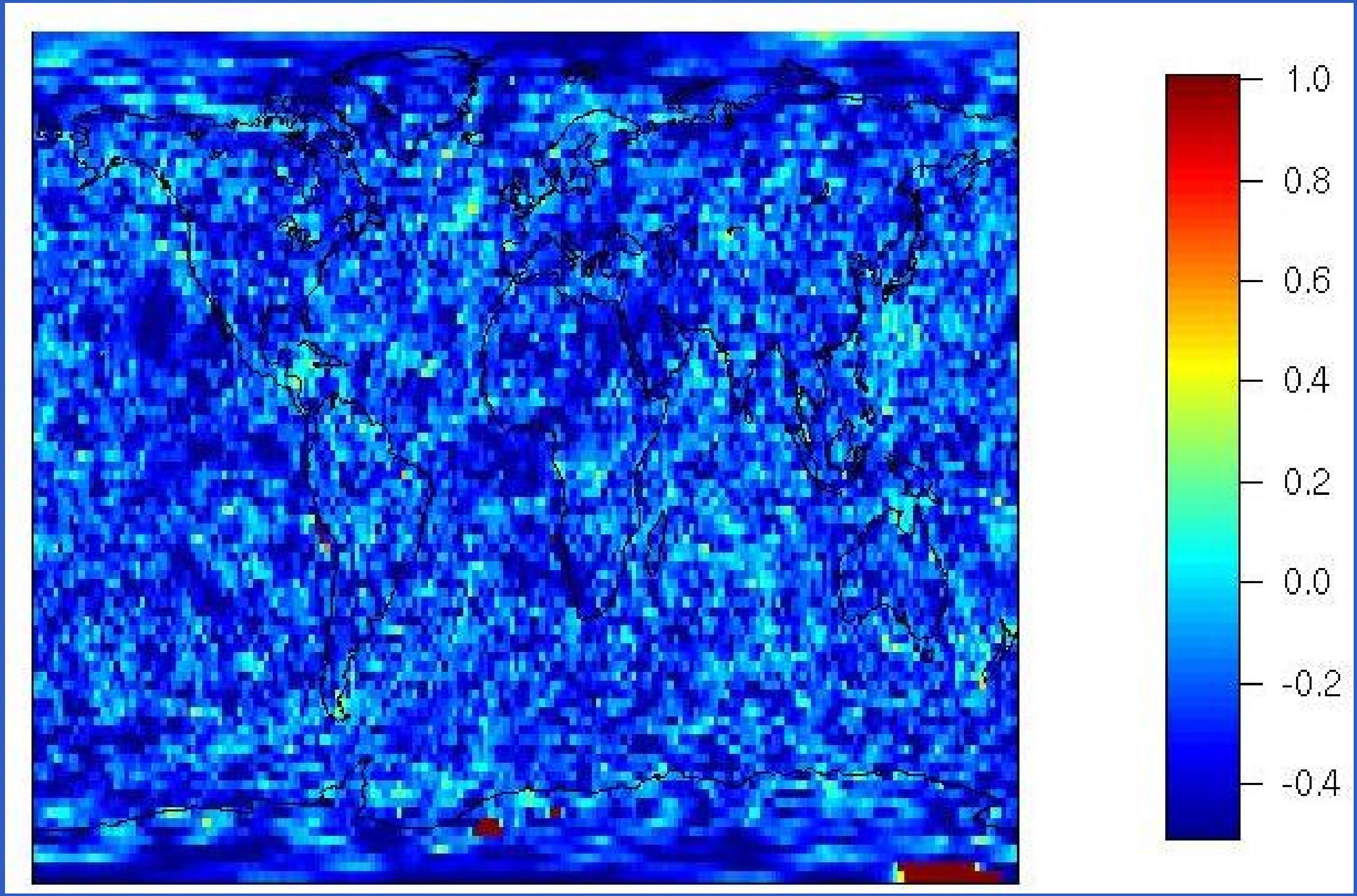
Shape Paramter - Max. CAPE



Shape Paramter - Max. Shear



Shape Paramter - Max. log*log



Next Steps

- Data is being translated into netcdf format to allow complete timeseries at single cells to be created. It will be easier to assemble the data needed to make peaks over threshold models.

- Smooth GEV parameters, then recalculate return levels.

- Allow parametric changes in GEV parameters to detect changes with respect to time.

- Fit point process models to account for both the magnitude and frequency of extreme cape, shear, or log(cape)log(shear).

- Compare reanalysis data to climate model data? (If climate model data have the same variables.)

Questions

- How should we take better advantage of the spatial aspect of the data?
- Since the data is to a degree a product of a statistical procedure, is it correct to think that it will have long tails?
- How might seasonality be addressed?
- Prior to 1980, data from the southern hemisphere is thought to be of poorer quality. How might this be addressed?

References

- Ventura V, CJ Paciorek, JS Risbey. (2004). Controlling the proportion of falsely rejected hypotheses when conducting multiple tests with climatological data, *J. of Climate* **17**:4343–4356.

- Craven, J. P., and H. E. Brooks, 2006: Baseline climatology of sounding derived parameters associated with deep, moist convection. *Nat. Wea. Digest, in press.*

- Brooks, Harold E., James W. Lee and Jeffrey P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data *Atmospheric Research*, Volumes 67-68, July-September 2003, Pages 73-94.

False Discovery Rate (FDR)

Benjamini and Hochberg (1995) FDR method (many unrealistic assumptions for climate data, such as independent data).

Control proportion, q, of falsely rejected null hypotheses relative to the total number of rejected hypotheses.

Let

$$T = \text{ Test Statistic}$$

$$t_i = \text{ observed value of } T \text{ at location } i;\ i = 1, \ldots, n$$

Large $\|T\|$ evidence against null hypothesis. The p-values are then defined as

$$p_i = \Pr\{\|T\| \geq \|t_i\| \mid\mid \text{null hypothesis is true}\}$$

False Discovery Rate (FDR)

Benjamini and Hochberg (1995) FDR method

Reject null hypothesis at all locations i for which the calculated p-values, p_i, are below a certain value. Specifically, $p_i \leq p_k$, where

$$k = \max_{i=0,\ldots,n} \{i\colon\ p_{i:n} \leq q\frac{i}{n}\}$$

with $p_{i:n}$ the i-th ordered p-value, and $p_{0:n} = 0$.

False Discovery Rate (FDR)

Ventura *et al*. (2004)

Modification of Benjamini and Hochberg method to allow for dependent data, and still have reasonable power. Same as above except:

$$k = \max_{i=0,\ldots,n} \{i : p_{i:n} \leq \frac{q'}{1-a}\frac{i}{n}\}$$

where a is the unknown proportion of true alternative hypotheses (Must be estimated!). Ventura *et al* suggest:

$$\hat{a} = I^{-1} \sum_{i=1}^{I} \max\left[0, \frac{\widehat{F}_p(x_i) - x_i}{1 - x_i}\right],$$

where $x_i = x_0(1-x_0)^{\frac{i-1}{I}}$, and I are regularly spaced intervals between x_0 and 1. Ventura *et al* suggest using $x_0 = 0.8$ and

$I = 20$ for optimal testing power. Also,

$$\widehat{F}_p(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{[0,x]}(p_i)$$